Session T2: Managing V&V

Session T2 leaders:

Co-Chairs:

Jamileh Soudah (Dept of Energy ASCI V&V lead)
Marty Pilch (Sandia National Laboratories)

T2 Materials in Foundations '02 proceedings:

Presentations (may contain back-up materials and notes)

Managing a Verification and Validation Program – The Government Perspective (26 slides) [T2B_barnes in both pdf and ppt formats]
Boots Barnes (JWARS Office)

Managing a Verification and Validation Program – The Contractor's Perspective (24 slides) [T2B_barnes in both pdf and ppt formats]
Mike Metz (IMC)

TOMAHAWK Simulation Management (24 slides) [T2B_barnes in both pdf and ppt formats]
Kem White (JHU/APL)

Managing the STORM V&V Program (35 slides) [T2B_barnes in both pdf and ppt formats]
Dave MacKay (IMC)

Panel Discussion of V&V Management:
Boots Barnes (JWARS Office)
Dave MacKay (IMC)
Mike Metz (IMC)
Kem White (JHU/APL)
Wendy Winner (Army Research Laboratory)

Participants in this session are listed at the end of the Discussion Synopsis.

Discussion Synopsis (to provide perspective on papers & briefings identified above).

Synopsis – because this session had only presentation, these synopses will help the briefings to be more understandable. Comments follow the synopses.

Kem White – Tomahawk Simulation Management

Tomahawk Simulation Management (TSM) Overview

The TSM has been established to address a problem of different organizations using simulations for different purposes – approach was ad hoc and not coordinated.

Organization with clear structure and responsibilities:

Includes Simulation Control Panel with SMEs make decisions on Accreditation of Tomahawk Simulation elements

Establish policies and procedures to define responsibilities and activities to be conducted within TSM

The Accreditation Process – the presenter went into detail about the accreditation process to include a description of the TSM's unique levels of accreditation:

> **Full Accreditation:** All requirements have been met.
>
> **Provisional Accreditation**: Performance has been demonstrated but not all accreditation requirements have been satisfied. (e.g., document preparation due to inadequate funding or sponsor accreditation timeline.)
>
> **Limited Accreditation:** Validation against real system data is incomplete although validation against all existing comparison data indicates correct performance and results consistent with other simulations.
>
> •Typically applies to an emerging development or when no real-world data exists.
> •Limitation on applicability is expressed on the certificate (e.g., Full accreditation for development flight test prediction).
>
> **Site Accreditation**: Applies when an accredited Tomahawk simulation is installed and used at an additional site.

The presenter described how the TSM accreditation process supported each test configuration.

The presenter discussed the measure of goodness of simulations to include a subjective SME judgment, view graph comparisons, and accreditation process.

Managing STORM VV&A by Dave MacKay

Presenter described how STORM fits into the Air Force Studies and Analysis (AFSAA) analysis process.

He also discussed how V&V is directly related to situational complexity and that V&V dominates at the mission level of simulation and accreditation dominates at the theater level campaign level of simulation due to the increased complexity and the difficulty of conducting V&V at that level.

The presenter described how the V&V process fits into the problem solving process between the M&S tools for analysis and the accreditation process as a basis for decision making.

He described US Air Force VV&A instructions and how accreditation supports the capability/usability/accuracy assessment of a simulation.

The STORM V&V process was then discussed in detail and described how it began in late 2001 with a goal of providing a decision package in 2004 for the addition of the STORM simulation to the Air Force Analysis Toolkit. The STORM V&V process uses the Analysis Toolkit criteria as its acceptability criteria in the process.

The STORM V&V process focuses on end-to-end traceability from requirements to code. It begins with the verified requirements, through the validated conceptual model, to the verified design. Code verification traces back to the verified design and results validation is back to the requirements and acceptability criteria.

Managing a V&V Program – a Government Perspective by Steven "Boots" Barnes

The presenter described the Joint Warfare System (JWARS) V&V effort that began in September 1997 and is still ongoing. He described how an integrated product/process team (IPT) was established early in the process to develop a tailored V&V process, a V&V Plan based on the process, and monitor and manage the execution of the V&V process in accordance with the plan.

He also described how the executive oversight committee and their steering committee managed the JWARS development and V&V processes. There was a description of the contractor's statement of work, how it was managed, and how the deliverables required were developed and delivered. He strongly endorsed the V&V Agent's use of a comprehensive JWARS V&V Database that provided a repository of development artifacts and all V&V related activities and reports.

The presenter described the $2.6M in funding provided for the first five years of the V&V effort and how that work supported a $70M development budget and how support was required from the JWARS Program Manager, the JWARS V&V Manager, and the Contract Officer's Representative. Also discussed was the thousands of hours provided by the JWARS User Sub Group members in each warfare area (SMEs) and the members of the IPT preparing for and attending meetings.

His recommendations to Government V&V Managers included:

- Make sure that your V&V Agent understands the simulation domain and the warfare domain (is not strictly a software tester)
- Form and use some type of oversight group/IPT to manage the process (include as many future users as possible)
- Use delivery or task orders to provide mechanism for changes without contract modification (one year or less for period of performance)

- Keep the V&V Agent independent of the developers (V&V is not QA or CM but they support each other)
- Combine the testing process with the V&V process when ever possible (reduces load on the developers)
- Provide a person in the sponsoring organization to coordinate V&V Agent and tester interaction with the developers
- Generate deliverable artifacts or database for future accreditation of work

Managing a V&V Program – a Contractor's Perspective by Michael Metz

This presentation mirrored the previous presentation from the Contractor's perspective. He began with how a joint venture team was formed to respond to the Government's RFP to meet all the required experience for the work. After contract award the V&V Agent began to work with the V&V IPT to develop the JWARS V&V process and plan.

He then described how the Operational Requirements Document (ORD) as approved by the Joint Requirements Oversight Council (JROC) included both a traceability and a V&V Key Performance Parameter (KPP) that were required. He then continued with a description of how he managed the Statement of Work (SOW) and deliverable performance.

The presenter also described how approximately 3.7 full-time equivalent people were provided each year with the $2.7M funds provided.

The presenter provided these recommendations:

- Educate your customer on what V&V is and what V&V isn't
- Use the DMSO managed *DoD VV&A RPG* plus any applicable Service guidance
- Propose the use of an oversight group/IPT to support the V&V process and include as many of the stakeholders in the simulation as possible, including future users
- Keep your focus on the current task or delivery order and let the sponsor know if you need changes
- Work hard to stay independent of the Developers to avoid being captured by their process and metrics
- Avoid attempts to group your effort with that of the Quality Assurance (QA) and Configuration Management (CM) processes
- Cooperate closely with the testers (Developer and Operational) and use their work to supplement yours
- If the Users are involved (a good idea) work closely with them
- Ask the Sponsor to provide a single point of contact for you within their organization and provide the same to them
- Tell the Developer about problems before you tell the Sponsor and work toward solutions
- Work hard to develop a relationship with the Developers to facilitate artifact collection and risk reduction
- Make sure you produce your deliverables on time and make them widely available

  - Risk reduction requires that feedback to Developers be in time to allow changes – don't wait until the final report
  - Don't assume that everyone on the Developer Team gets your products
  - Use the Developers Change Request Process
- Save and document everything

.

Topics and Issues Raised and Discussed in the Session

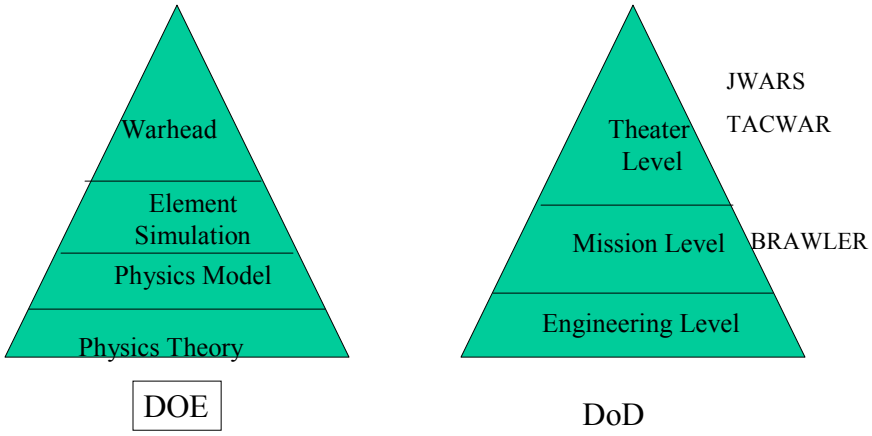
Suggestions:

Presenters should avoid use of domain specific acronyms that people in other communities don't know. In future sessions like this, an effort should be made to include presentations from M&S activities outside of DoD.

Issues and Topics

- Concerns that cost and schedule drive V&V efforts instead of the requirements for V&V.
- The importance of Referents in the domain functionality validation and results validation processes.
- What to do when there are no Referents (including SME knowledge) for validation.
- There was a discussion about the use of Sensitivity Analysis in VV&A.
- There was a discussion about how the problems of people in different areas (varying level of simulations from engineering level through campaign level in DoD)

Modeling Hierarchy Comparison



Q: What is an "audit of technical data?"
A: An expert review, sometimes called "face validation."

Q: Do you use any methods that include quantitative comparisons of real world and simulation results graphs with a requirement that the simulation and the test must be within 10% of each other?
A: No.

Q: Do the Navy and COMOPTEVFOR accept Tomahawk Simulation Management (TSM) accreditation?
A. Yes but for limited use. TMS uses a unique description of states of accreditation.

Q: Is "Value Added" measurable?
A: Maybe by milestones – I think our biggest value added has been finding errors in software.

Q: Who looks at the Conceptual Model? How are the requirements vetted? Is all this being reviewed by an inside group?
A: The STORM IPRs are attended by many of the future users and they are the reviewers of the conceptual model and the requirements. AFSAA doesn't do it alone – they use the V&V Agent and work with the future users.

Q: How do you plan the duration of the V&V activities?
A: The overall process is cost and schedule driven. We do what we can afford to do within the plan and the schedule.

Q: V&V efforts that get mounted appear to be\ forced into a particular cost and schedule regime. Are we really doing V&V or are we just checking the box? Do you feel that budgeting limitations prevent you from doing what you would like to do?
A: V&V Agents are under the gun and always wonder if they're going to have time and resources to do everything that needs to get done under the current schedule. However, even if V&V Agents don't do everything or do it to the depth they feel they should it is still a value added process – not window dressing. Although cost and schedule drive it we have to understand the risk and work to mitigate it as best we can by identifying the high risk areas and doing them first.

Q: For V&V – how's your role different than Quality Assurance (QA)? Please give some examples.
A: Here's an example – you may ask someone to model an existing weapons system. It may require the conduct of an experimental program. The risk – to the users – of not properly modeling the system is great. In my opinion QA would support the software development activities but would not address the simulation capability issues that are the focus of the risk.

- QA checks the software. If QA people are testers they're conducting unit verification steps. If they're not testers its more like an administrative activity.
- QA in M&S is very new. An analogy from software QA is that QA is a milestone in the process and the rules are common to all. V&V is mostly related to the specific function to be modeled and simulated – the rules aren't common to all.

Q: How about data V&V? Where does it occur in the process?
A: We try to V&V data inputs throughout the process. Data V&V occurs from conceptual model validation through results validation.

Q: Do we have to verify (and validate) models only once? Or for every intended use?
A: Prior V&V or accreditation results will be used and accepted where possible. If it is determined that more V&V information is required for the new intended use it will be required. The user must do a risk assessment. How much use is there in using the simulation as it is for my different intended use. If the risk is unacceptable or unknown additional V&V will probably be required.

Q: Is V&V a one-time activity if the intended use does not change?
A: No. New verification may be required by code changes and subsequent validation events will probably be required.

Q: I don't have results for comparison to use as a referent in my V&V effort. What should I do? Who should bless my results?
A: If there is no test or real world referent you are forced to fall back on the use of SME knowledge.

Comments:

- It appears that each of the V&V programs described today uses the Integrated Product/Process Team (IPT) approach.
- Since we don't know what our results should be we are using sensitivity analysis for things we don't or can't know to see if the unknown behavior will affect our results.
- We do V&V because if we don't our customers won't come to us and use our models. We use it to make our models credible to the users.

Q: What's the definition of Results Validation.
A: The comparison of the output of the simulation to a referent. The real world is the best referent but is rarely available. Other choices are test results, other M&S results, or SME knowledge.

Comments:

- There was discussion of the use of statistical hypothesis in results validation.
- Accreditation is a political statement. There is some blanket accreditation in the Army (probably Class Accreditation). In DoD some people will take the risk to accredit without additional V&V. I suggest that you look for old data and decide how much risk you're willing to take in accrediting a simulation.

Q: The issue of adequacy has not been addressed for simulations. Do you use expert judgment to determine if a simulation is adequate?
A: Depends on the type of model. An engineering level model can probably be compared to the real world piece of equipment. SMEs are frequently used for campaign level analysis validation. Other times the engineering level or mission level activities that can be validated are used as referents to validate campaign level models.

Comments:

- Suggest that you validate to the best referent you can find.
- If you're using stochastic simulations suggest you use some quantitative tests for your random number generators to ensure that they're as random as possible given the limitations of random number generators.

T2 Session Participants (19)

First Name	Last Name	Organization
"Boots"	Barnes	JWARS Office
Dave	Bort	JHU/APL
Cam	Catts	Lockheed Martin
Mike	Dennis	JHU/APL
Scott	Doebling	Lawrence Livermore National Lab
Marti	Hoppus	JASA/NAWCWD
Roger	Logan	Lawrence Livermore National Lab
David	MacKay	AFSAA/SAAP
Michael	Metz	Innovative Management Concepts, Inc
Orhun	Molyer	DND/CF SECO
Cynthia	Nitta	U.C./Lawrence Livermore National Lab
James (Chuck)	Segrest	Lockheed martin
Seth	Shepherd	USAF 412TW
Jamileh	Soudah	DOE/NNSA
Joseph	Tasto	Immersion Medical
Robert	Thomas	Sandia National Laboratories
Donald	White	Johns Hopkins University
Wendy	Winner	USARL/SLAD
William	Yeakel	Orsa Corporation